基于图生成过程的跨领域推荐 *

蔡瑞初†，吴逢竹，李梓健

(广东工业大学 计算机学院，广州 510006)

摘　要：推荐系统在各方各面得到充分的应用，时刻影响着日常生活。要训练出一个良好的推荐系统往往需要大量的“用户-商品”交互数据，但是实际情况下获得的数据往往是十分稀疏的，这往往会使得训练出来的模型过拟合，最后难以获得理想的推荐效果。为了解决这个问题，跨领域推荐系统应运而生。目前大部分的跨领域推荐系统工作都是借鉴传统领域自适应的方法，使用基于特征对齐或者对抗学习的思想将领域不变用户兴趣从有丰富数据的源域迁移到稀疏的目标域上，例如豆瓣电影迁移到豆瓣图书。但是由于不同推荐平台的网络结构有所不同，现有的方法暴力提取的领域不变的语义信息容易和结构信息耦合，导致错配现象。而且，现有的方法忽略了图数据的本身存在的噪声，导致实验效果进一步受到了影响。为了解决这个问题，首先引入了图数据的因果数据生成过程，通过领域特征隐变量和语义特征隐变量、噪声隐变量解耦出来，通过使用每个节点的语义隐变量进行推荐，从而获得领域不变的推荐效果。本文在多个公开数据集上验证了本文的方法，并取得了目前最好的实验效果。

关键词：解耦；图神经网络；领域自适应；推荐系统；因果数据生成过程

中图分类号：TP399　　doi: 10.19734/j.issn.1001-3695.2022.01.0015

Cross-domain recommendation under graph generation process

Cai Ruichu†, Wu Fengzhu, Li Zijian

(*School of Computer Science*, *Guangdong University of Technology*, *Guangzhou* 510006, *China*)

Abstract: Recommendation systems are widely used everywhere and have a great influence on daily life. Aiming to train an ideal recommendation system, a massive of ‘use-item’ interactive pairs should be provided. However, the dataset obtained is usually sparse, which might result in an over-fitting model and be hard to obtain the satisfying performance. In order to address this problem, the cross-domain recommendation is raised. Most of the existing methods on cross-domain recommendation systems borrow the ideas of the conventional unsupervised domain adaptation, which employ the feature alignment or adversarial training methods. Hence they can transfer the domain-invariant interests of users from the source to the target domains, e. g. , from the Douban Movies to the Douban Books. However, since the network structures vary with different recommendation platforms, the existing methods on cross-domain recommendation systems straightforwardly extract the domain-invariant representation may entangle the structure information, which may result in the false alignment phenomenon. Furthermore, the previous efforts ignore the noise information behind the graph data, which further degenerate the experimental performance. In order to address the aforementioned problems, this paper brings the causal generation process of graph data into the cross-domain recommendation systems, this paper use the semantic latent variables of each node to calculate the relationships between users and items via disentangling the semantic latent variables, domain latent variables and noise latent variables. Experiments studies testify that the proposed method yields state-of-the-art performance on several cross-domain recommendation system benchmark datasets.

Key words: disentanglement; graph neural networks; domain adaptation; recommendation system; causal generation process

0　引言

近年来，随着深度学习技术的迅速发展，深度学习在各行各业遍地开花，其中的一个应用是推荐系统。推荐系统是为推荐商品而设计的特殊软件程序，其目标是从海量的候选商品集中预测用户未来最可能交互(点击、评价、购买等)的商品。协同过滤[1](Collaborative filtering, CF)方法是推荐系统的基石，它基于用户过去的交互(如产品评分)，对用户对商品的偏好进行建模。矩阵分解[2](Matrix factorization, MF)是一类 CF 方法，它通过分解用户和物品的交互矩阵来学习用户和物品的潜在因素。神经协同过滤是另一类 CF 方法，利用神经网络学习复杂的用户-物品交互函数[3]。然而，传统的矩阵分解和神经网络协同过滤都存在冷启动和数据稀疏问题。

一个有效的解决方法是转移相关领域的知识，跨领域推荐[4]技术解决了这类问题。在现实生活中，用户通常会参与几个系统来获得不同的信息服务。例如，用户在在线视频网站看电影的同时，从知乎或豆瓣等网站浏览书籍评价。本文可以通过跨领域学习来提高目标服务(或所有服务)的推荐性能。如图 1(a)所示，相同用户在不同平台的行为是相似的，喜欢看西游记书籍的人，也会喜欢看西游记的电影。按照上面的例子，本文可以用两个矩阵分别表示视频与图书的浏览反馈，通常这两个矩阵是高度稀疏的，同时学习它们有益于缓解数据稀疏。集合矩阵分解(Collective MF, CMF)[5]是这类方法的代表，它通过共享用户潜在因素来共同分解这两个矩

收稿日期：2022-01-06；**修回日期：**2022-03-03　**基金项目：**国家自然科学基金资助项目(61876043)；国家优秀青年科学基金资助项目(6212200101)；广州市科技计划项目(201902010058)

作者简介：蔡瑞初(1983-)，男(通信作者)，浙江温州人，教授、博导，主要研究方向为因果关系、机器学习(cairuichu@gmail.com)；吴逢竹(1997-)，男，广东佛山人，硕士研究生，主要研究方向为图表示学习、迁移学习；李梓健(1994-)，男，广东广州人，博士研究生，主要研究方向为因果性学习、迁移学习.

阵，将目标域的 CF 和辅助域的 CF 相结合实现知识转移。然而，CMF 是一个浅层模型，难以学习复杂的用户-物品交互函数。另一类方法是利用图神经网络[6,7](Graph Neural Network, GNN)，建模交互图的高阶连通性信息。图神经网络在嵌入空间中构建信息流，利用图卷积层传播节点嵌入，并聚合交互过的物品(或用户)的嵌入来强化用户(或商品)的嵌入。通过叠加多个图卷积层，本文可以在节点嵌入中捕获高阶连接中的协同信号，以此缓解数据稀疏的问题。

虽然现有的方法一定程度上缓解了跨领域推荐问题，但是这些方法大多是简单地借鉴非图数据的无监督领域自适应方法，这些方法旨在提取领域不变的信息，但是暴力地采用这些方法提取出来的信息会使得领域特有的结构信息和领域不变的用户兴趣信息耦合，使得不同用户信息迁移错误。如图 1(b)所示，现有方法提取不变信息的同时，引入了领域特有的结构信息，导致错配现象。(用户迁移错误)，如果领域不变表达中同时语义信息和结构信息耦合，那么会出现不同用户对齐的现象(例如男孩迁移到女孩)，最终导致不理想的推荐效果。为了解决这个问题，一个直观的解决方法是将语义信息和结构信息解耦出来，然后在推荐的时候只使用语义信息，从而可以避免结构信息的干扰。如图 1(c)所示，提取网络特征的时候通过解耦出语义隐变量，避免领域特有的结构特征的影响，从而能够进行有效的迁移,可以通过舍弃不同推荐平台数据中的结构信息，使得相同用户在不同推荐平台上的行为得到对齐。

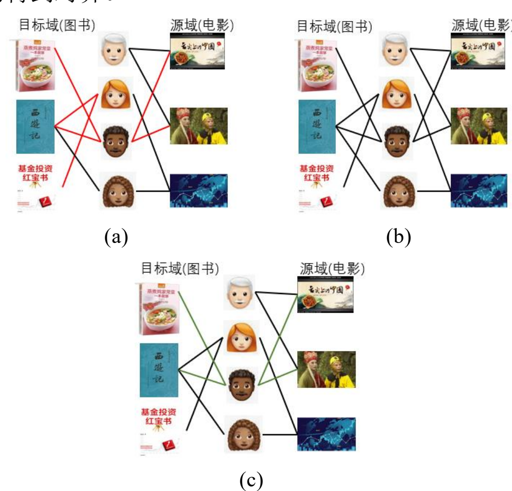


图 1　一个简单的跨领域推荐系统例子

Fig. 1　A simple example of cross-domain recommendation system

为了将语义信息和结构信息解耦出来，本文假设了如图 2 所示的针对图数据的因果数据生成过程。如图 2 所示，本文假设 d 表示不同领域(如不同的推荐平台)，y 表示不同用户/商品，A_o 表示不同推荐平台的噪声。$\mathbf{h}_u^{sem}$，$\mathbf{h}_u^{dom}$，$\mathbf{h}_u^{noi}$ 分别表示结构隐变量，语义隐变量和噪声隐变量，本文观测到的图数据 $\mathcal{G}$ 由这三个隐变量同时生成。

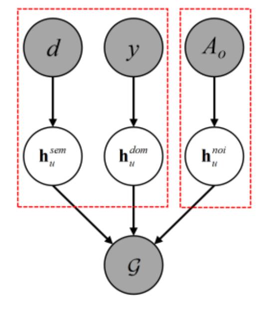


图 2　基于图数据的因果数据生成过程

Fig. 2　Causal data generation process based on graph data

基于上述针对图数据的数据生成过程，并提取的领域不变的用户兴趣表示，本文提出了一种基于图生成过程的跨领域推荐方法。本文假设用户-物品交互图的生成过程由两种独立的潜在变量控制，即语义潜在变量 $\mathbf{h}_u^{sem}$ 和域潜在变量 $\mathbf{h}_u^{dom}$。通过对这两个潜在变量的重构和分解，本文可以很容易地在 $\mathbf{h}_u^{sem}$ 的基础上进行下游的推荐任务。在技术上，本文采用图自编码器[28]来重建这些潜在变量，并利用得分预测器和领域分类器来分解语义和领域潜在变量。大量的实验研究表明，本文的方法在多个公开数据集上的性能优于目前的跨领域推荐方法。

本文提出基于图生成过程的跨领域推荐系统，其研究意义如下：首先，跨领域推荐能同时利用多个数据源，解决推荐系统中经常面临的冷启动问题和数据稀疏问题；其次，目前的跨领域推荐研究大都是针对传统的深度神经网络模型，对更复杂的图网络探讨较少，特别是如何解决领域特有的结构信息和领域不变的用户兴趣信息的耦合，还存在很多这方面的研究空白；最后，本文首次在跨领域推荐中引入了因果知识，通过因果数据生成过程提升了模型的稳定性和鲁棒性。

1　相关工作

1.1　推荐系统

推荐系统旨在从用户的历史行为中了解他们对未知物品的偏好。现有的推荐系统可以分为基于内容的推荐(Content-based Recommendation)[8]与协同过滤(Collaborative Filtering，CF)[1,9~12]两大类。基于内容的推荐是基于用户画像和物品描述之间的匹配。当缺少关于物品的内容描述时，基于内容推荐的方法难以为每个用户构建个性化的画像。协同过滤通过基于用户与物品交互行为预测用户偏好，与物品的内容无关，以此缓解内容信息稀缺的问题。潜在因子模型主要基于矩阵分解(Matrix Factorization，MF)[2]学习用户和物品的特征向量，且具有概率解释。分解机(Factorization Machines，FM)[13]是广义的矩阵分解，能有效建模多种特征之间的交互。神经协同过滤(Neural Collaborative Filtering, NCF)[3]把深度神经网络与矩阵分解相结合，学习非线性的特征表示。然而，真实世界的用户-物品交互数据是高度稀疏的，导致协同过滤模型在训练时受到数据稀疏性的影响。

跨领域推荐(Cross-domain Recommendation)[4,14,15]的目的是利用源领域充足的数据，缓解目标领域的用户行为稀缺的问题。有一类把矩阵分解应用于每个域的跨领域推荐方法，例如集合矩阵分解 (Collective MF，CMF)[16]。上述浅层的方法，在学习高度非线性的复杂用户-物品交互关系方面存在困难。此外，DARec[17]借鉴迁移学习中的域自适应技术进行跨领域推荐，TMCDR[18]借助模型无关的元学习解决跨领域的用户冷启动问题。

本文通过引入图数据的因果数据生成过程，将语义隐变量和领域隐变量同时解耦出来，从而避免了由结构引起的用户信息错配，获得更好的迁移效果。

1.2　图神经网络

另一个相关的研究领域是基于用户-物品二部图结构进行推荐，近年来出现的图神经网络(Graph Neural Networks, GNN)[6,7]广泛应用于图结构数据的建模，尤其是高阶邻居，以指导节点表示的学习。早期的研究定义了谱域上的图卷积，如基于拉普拉斯特征分解和 Chebyshev 多项式的图卷积[19]，但这类方法需要计算拉普拉斯矩阵的逆，计算复杂度较高。随后，图卷积网络(Graph Convolutional Networks，GCN)[6]和 GraphSAGE[7]在空间域中重新定义了基于消息传递的图卷积，即通过聚合邻居的表示来活得目标节点的表示。由于其可解释性和计算高效，它们迅速成为一种流行的图神经网络范式并被广泛应用[20,21]。得益于图卷积的强大表达能力，大量基

于图网络的推荐模型被提出[22~29]，如 GC-MC[27]，NGCF[28]和 LightGCN[29]使 GCN 适用于用户-物品交互图的建模，捕获目标节点高阶邻居的协同过滤信号进行推荐。在最近的基于图网络推荐的研究中，SGL[30]探索了对用户-物品图的自监督学习，以提高 GCN 推荐的多样性和稳健性；SGCN[31]利用图的稀疏性和低层次的结构特性，在 GCN 的每一层都附加了一个可训练的随机二进制掩码，以此提高 GCN 的推荐性能。

还有一些研究把跨领域推荐与图神经网络相结合，例如 PPGN[32]构建了一个跨领域偏好矩阵来对不同领域的交互作为一个整体进行建模，以更好的捕捉用户偏好在交互图中的传播方式；CD-GNN[33]在由社交网络与用户浏览记录组成的异构图上学习用户的表示，然后提取的领域不变特征用于推荐。本文将图神经网络应用到图生成过程中建模，利用图神经网络的点聚合能力，学习出不同的隐变量。

2　基于图生成过程的跨领域推荐

2.1　问题定义

给定两个领域，一个源领域(例如，电影推荐)和一个目标领域(例如，书籍推荐)，两个领域重合的用户集合记为 $\mathcal{U}$，源领域和目标领域的物品分别记为 $\mathcal{V}_s$ 和 $\mathcal{V}_t$。每个领域的任务都是隐式反馈的协同过滤问题。对于源领域，本文用一个无向图 $\mathcal{G}_s$ 来描述用户与物品的交互，图有 $\mathcal{U}$ 和 $\mathcal{V}_s$ 两类节点，如果用户 $u\in\mathcal{U}$ 与物品 $v\in\mathcal{V}_s$ 存在交互，则在 u 与 v 之间置一条边。对于目标领域，以同样的方法构建 $\mathcal{G}_t$。本文的符号表如表 1 所示。

在商品推荐任务里，每个用户只对前 N 个物品感兴趣，而物品根据他们的预测评分排序，公式如下：

$$\hat{y}_{uv}=f(u,v\mid\Theta) \tag{1}$$

其中，f 是交互函数，Θ 是模型参数。对于矩阵分解，交互函数是固定的点积，公式如下：

$$\hat{y}_{uv}=\boldsymbol{e}_u^T\boldsymbol{e}_v \tag{2}$$

其中，$\boldsymbol{e}_u\in\mathbb{R}^{d\times1}$ 和 $\boldsymbol{e}_v\in\mathbb{R}^{d\times d}$ 是用户 u 与物品 v 的维嵌入向量，那么模型参数就是由嵌入向量组成的嵌入矩阵 $\boldsymbol{U}\in\mathbb{R}^{d\times|\mathcal{U}|}$ 和 $\boldsymbol{V}\in\mathbb{R}^{d\times|\mathcal{V}|}$。在跨领域推荐任务中，本文利用两个领域的数据协同训练模型，通过知识迁移来提高模型表现。

表 1　符号表

Tab. 1　Notation table

符号	描述
u, v_s, v_t	用户节点、源领域物品节点与目标领域物品节点
$\mathcal{U},\mathcal{V}_s,\mathcal{V}_t$	用户集合、源领域物品集合与目标领域物品集合
$\mathcal{G}_s,\mathcal{G}_t$	源领域的交互图与目标领域的交互图
$\mathcal{N}_u^{(s)},\mathcal{N}_u^{(t)},\mathcal{N}_v$	u 在 $\mathcal{G}_s$ 上的邻居、u 在 $\mathcal{G}_t$ 上的邻居和 v 的邻居
$\boldsymbol{e}_u^{(l)},\boldsymbol{e}_v^{(l)}$	第 l 层图卷积输出的用户表示与物品表示
$\boldsymbol{h}_u,\boldsymbol{h}_v$	图卷积模块提取的用户表示与物品表示
$\boldsymbol{h}_u^{sem},\boldsymbol{h}_u^{dom},\boldsymbol{h}_u^{noi}$	用户的语义隐变量、领域隐变量、噪声隐变量
$y_{uv},\hat{y}_{uv}$	u 与 v 是否交互的真实值与预测值
$d_u,\hat{d}_u$	领域分类器输出的用户 u 领域的预测值

在训练阶段的每个批次，本文从 $\mathcal{U}$ 中均匀采样一批用户；对于每个用户，从源领域的训练集中随机采样一个交互过的物品 v_S 作为正样本，再从 $\mathcal{V}_s$ 中均匀采样一个未交互过的物品作为负样本 v_S'；对目标领也进行同样的采样。

2.2　模型设计

本文从生成模型的角度去考虑跨领域推荐问题，模型的实现参考了图自编码器[28]，整体上由编码器与解码器组成，模型框架如图 3 所示。模型由嵌入层、图卷积层、解耦模块、领域判别器、源领域得分预测器、目标领域得分预测器和重构模块这 7 个部分组成，本文在后续章节中详细介绍每一个模块。

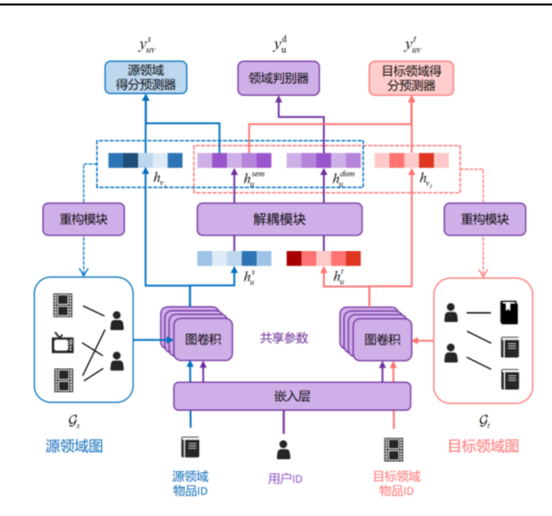


图 3　模型结构

Fig. 3　Model structure

2.3　嵌入层与图卷积层

嵌入层负责把节点 ID 转换为对应的嵌入向量，具体地，本文输入用户 u 的 ID，从跨领域共享的用户嵌入矩阵 $\boldsymbol{U}$ 中提取其嵌入向量 $\boldsymbol{e}_u^{(0)}$；输入源领域物品 v_S 的 ID，从源领域物品的嵌入矩阵 $\boldsymbol{V}_s$ 提取其嵌入向量 $\boldsymbol{e}_{v_s}^{(0)}$，对于目标领域，以相同方式从 $\boldsymbol{V}_t$ 提取 $\boldsymbol{e}_{v_t}^{(0)}$。所以，嵌入层包含 $\boldsymbol{U},\boldsymbol{V}_s,\boldsymbol{V}_t$ 这三组训练参数。由于源领域与目标领域是对称的，为了简化公式，下文用 v 指代 v_S 与 v_t。

图卷积层负责从二部图上学习节点表示，它接收 $\mathcal{G}_s$ 或 $\mathcal{G}_t$ 以及目标节点作为输入，以目标节点为中心，逐层聚合邻域节点的表示，公式如下：

$$\boldsymbol{e}_u^{(l+1)}=\sum_{v\in\mathcal{N}_u}\frac{1}{\sqrt{|\mathcal{N}_u|}\sqrt{|\mathcal{N}_v|}}\boldsymbol{W}_{uv}^{(l)}[\boldsymbol{e}_u^{(l)}\parallel\boldsymbol{e}_v^{(l)}] \tag{3}$$

$$\boldsymbol{e}_v^{(l+1)}=\sum_{u\in\mathcal{N}_v}\frac{1}{\sqrt{|\mathcal{N}_u|}\sqrt{|\mathcal{N}_v|}}\boldsymbol{W}_{vu}^{(l)}[\boldsymbol{e}_v^{(l)}\parallel\boldsymbol{e}_u^{(l)}] \tag{4}$$

$$\boldsymbol{h}_u=\parallel_{l=0}^{L}\boldsymbol{e}_u^{(l)} \tag{5}$$

$$\boldsymbol{h}_v=\parallel_{l=0}^{L}\boldsymbol{e}_v^{(l)} \tag{6}$$

其中，$\parallel$表示向量拼接操作，$\mathcal{N}_u$ 表示节点 u 的邻居。$\boldsymbol{W}_{uv}^{(l)}\in\mathbb{R}^{d\times2d}$ 和 $\boldsymbol{W}_{vu}^{(l)}\in\mathbb{R}^{d\times2d}$ 是第 l 层图卷积网络的可训练参数，本文用不同的权重矩阵建模以不同节点为中心的聚合函数，希望网络能学到用户与物品之间非对称的关系。本文叠加多层图卷积，每层图卷积的输入是上一层的输出。最后，把每一层的输出拼接获得用户节点表示 $\boldsymbol{h}_u$ 与物品节点表示 $\boldsymbol{h}_v$。

2.4　解耦模块

为了获得解耦的用户兴趣表示，本文进一步把 $\boldsymbol{h}_u$ 分别输入隐变量解耦模块，获得用户语义向量 $\boldsymbol{h}_u^{sem}$，用户领域向量 $\boldsymbol{h}_u^{dom}$ 和噪声向量 $\boldsymbol{h}_u^{noi}$。$\boldsymbol{h}_u^{sem}$ 表示用户跨领域的兴趣，可用于计算与物品的得分；而 $\boldsymbol{h}_u^{dom}$ 仅包含与领域有关的信息，用于输入领域判别器预测样本的领域；$\boldsymbol{h}_u^{noi}$ 包含与领域和语义都无关的噪声。本文用 MLP 网络作为语义编码器与领域编码器，其公式如下：

$$\boldsymbol{h}_u^{sem}=\mathrm{MLP}(\boldsymbol{h}_u;\theta_{sem}) \tag{7}$$

$$\boldsymbol{h}_u^{dom}=\mathrm{MLP}(\boldsymbol{h}_u;\theta_{dom}) \tag{8}$$

$$\boldsymbol{h}_u^{noi}=\mathrm{MLP}(\boldsymbol{h}_u;\theta_{\mathrm{noi}}) \tag{9}$$

其中，θ_{sem}，θ_{dom} 和 θ_{noi} 为 MLP 网络的可训练参数。

2.5　领域判别器、源领域得分预测器和目标领域得分预测器

由于推荐任务的语义信息是用户对物品的评分，本文把 $\boldsymbol{h}_u^{sem}$ 输入评分预测器，以评分作为监督新号，从 $\boldsymbol{h}_u$ 中提取有关语义的信息。考虑到本文没有用到除物品 ID 外的其他特征，本文采用与物品表示的内积作为预测的评分，并采用 BPR[29]损失函数进行训练，公式如下：

$$\hat{y}_{uv}=\langle\boldsymbol{h}_u^{sem},\boldsymbol{h}_v\rangle \tag{10}$$

$$L_y = \sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{N}_u}\sum_{v'\notin\mathcal{N}_u} -\ln\sigma(\hat{y}_{uv}-\hat{y}_{uv'}) \tag{11}$$

其中，$\sigma(\cdot)$ 是 Sigmoid 函数，v'是从物品集合中均匀采用的未与用户交互过的物品，即负样本。BPR(Bayesian Personalized Ranking, BPR)[34]损失函数并不直接优化预测的得分，而是最大化正样本与负样本得分的差值，使模型学到“相比于 v'，u 更喜欢 v”这个偏好。

对应的，本文把用户领域向量 $\boldsymbol{h}_u^{dom}$ 输入领域判别器，以领域标签 d_u 作为输入监督信号，若物品来自源领域，d_u=0，否则 d_u=1。领域判别器采用一个 MLP 网络实现，并采用交叉熵作为损失函数，公式如下：

$$\hat{d}_u = \text{MLP}(\boldsymbol{h}_u^{dom};\theta_{cls}) \tag{12}$$

$$L_d = \sum_{u\in\mathcal{U}} d_u\log(\hat{d}_u)+(1-d_u)\log(1-\hat{d}_u) \tag{13}$$

其中，θ_{cls} 是领域判别器的可训练参数。

2.6　重构模块

为了重构输入的图结构，从而最大程度的保留图的结构信息，本文采用了图自编码器[28]的做法，用两个节点表示的内积作为边存在的概率，以此恢复输入图的邻接矩阵。本文利用一个 MLP 重新聚合了解耦后的 $\boldsymbol{h}_u^{sem}$、$\boldsymbol{h}_u^{dom}$ 和 $\boldsymbol{h}_u^{noi}$，以此作为用于重构的用户表示 $\boldsymbol{h}_u^{r}$。考虑到完整的邻接矩阵过于庞大，本文仅恢复用户节点所在网络的局部结构，并采用 word2vec 的负采样技术。重构模块的公式如下：

$$\boldsymbol{h}_u^{r} = \text{MLP}(\boldsymbol{h}_u^{sem}\,\|\,\boldsymbol{h}_u^{dom}\,\|\,\boldsymbol{h}_u^{noi};\theta_r) \tag{14}$$

$$L_r = \sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{N}_u}\log(\sigma(\langle\boldsymbol{h}_u^r,\boldsymbol{h}_v\rangle)) + \sum_{v'\in\mathcal{V}-\mathcal{N}_u}\log(1-\sigma(\langle\boldsymbol{h}_u^r,\boldsymbol{h}_{v'}\rangle)) \tag{15}$$

其中 θ_r 是可训练参数，源领域与目标领域的重构模块共享参数。

2.7　模型训练与测试

在训练阶段，本文分别计算源领域与目标领域的得分损失 L_y^s 和 L_y^t，领域分类器的损失 L_d，源领域与目标领域的重构损失 L_r^s 和 L_r^t，以及对所有模型参数 Θ 的 L2 正则化项 $\|\Theta\|_2^2$。模型的总损失函数如下：

$$L = L_y^s + L_y^t + \alpha L_d + \beta(L_r^s + L_r^t) + \lambda\|\Theta\|_2^2 \tag{16}$$

其中，α 是控制领域损失权重的超参数，默认值为 1.0，β 是控制重构损失的超参数，默认值为 1e-3，λ 是控制 L2 正则化权重的超参数，默认值为 1e-5，。本文采用 Adam 优化器优化总损失函数并更新模型参数。

在测试阶段，由于本文只关注目标领域的用户-物品得分，本文仅计算图卷积层的式(3)~(6)与语义编码器的式(7)的输出，然后通过式(10)获得目标领域的预测得分，其他与目标领域任务无关的模块不需要计算。

3　实验设计与分析

3.1　数据集介绍

Amazon 2018 评论数据集(https://nijianmo.github.io/amazon/index.html)该数据集由 Amazon 电商平台上 1996 年 6 月至 2018 年 10 月的共 2.3 亿条商品评论构成，每条评论包含用户 ID、商品 ID、时间戳、评论文本、用户评分以及商品的元数据(类别、品牌、描述、图像特征等)。Amazon 评论数据集为跨领域推荐提供了一个测试平台，具体任务是为用户推荐他可能会交互的商品。为了把显示反馈的评分转换为隐式反馈的形式，本文仅保留评分大于等于 3 的评论作为样本。然后，本文根据商品类别划分评论数据集，每个类别商品的评论作为一个领域，选择其中数据较为丰富的 CDs、Digital Music、Movies 和 Books 分别作为源领域和目标领域，各个数据集的统计信息见表 2，其中“ → ”两边的数据集分别代表源领域和目标领域。

由表 2 可知，第一组数据集的源领域 CDs 与目标领域 Digital Music 较为相近，物品语义与交互网络的稀疏程度都相似；而第二组中，目标领域 Books 的稀疏程度与物品数量都远超源领域 Movies，因此比第一组数据集更有挑战性。关于数据集的划分，本文采用留一法(Leave-one-out)[3]，首先把每个用户交互过的商品按时间戳升序排序，选择最后一个商品作为训练集，倒数第二个商品作为验证集，剩余的商品作为训练集。为了保证数据质量，本文仅保留交互次数大于 3 的用户和商品，且保证每个用户在源域和目标域都有交互记录。最后，关于用户与商品的特征选择，本文关注的是结构信息在跨领域推荐的作用，为了公平比较各种方法，输入特征仅包含用户 ID 与物品 ID。

表 2　Amazon 数据集的统计信息

Tab. 2　Amazon dataset statistics

数据集	用户数量	物品数量	交互数量	稀疏度(%)
CDs	20,084	112,532	504,679	0.0223
Digital Music		30,393	148,562	0.0243
Movies	99,587	54,023	1,439,951	0.0268
Books		504,017	2,718,225	0.0054

3.2　实验设置

对于商品推荐任务，学术界普遍采用的留一法(Leave-one-out)进行评估，即保留每个用户最新的交互用于测试，第二新的交互作为验证集，用于调整超参数，剩余的交互作为训练集。在测试时，对于每个用户，模型为所有的候选物品打分并按照得分降序排序，然后返回得分最高的前 N 个物品作为该用户的推荐结果，这种任务被称为 Top-N 推荐。由于同时考虑所有候选物品的计算量过大，本文参考[3]的策略，随机采样 500 个没有与用户交互过的物品作为负样本，然后评估模型把正样本(测试物品)排在负样本前的能力。典型的 Top-N 评估指标有 HR(Hit Rate)和 NDCG(Normalized Discounted Cumulative Gain)。HR 衡量测试物品是否存在于模型输出的 Top-N 列表中，其定义如下：

$$HR@\text{N} = \frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}}\delta(r_u\le N) \tag{17}$$

其中 r_u 是用户 u 的测试物品在返回的 Top-N 列表中的排序，$\delta(\cdot)$ 代表指示函数(indicator function)。NDCG 在 HR 的基础上考虑了测试物品的具体排名，其定义如下：

$$NDCG@\text{N} = \frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}}\frac{\log 2}{\log(r_u+1)} \tag{18}$$

测试物品在 Top-N 列表中的排名越靠前，则权重越大。对于 HR 和 NDCG，指标越高代表模型表现越好，本文给出 N=[5,10]的实验结果。

本文基于 PyTorch (https://pytorch.org/)和 Deep Graph Library (https://github.com/dmlc/dgl)框架实现了所有的基线模型，并提供可复现的 Python 代码(https://github.com/rynewu224/CrossDomainGNNRec)。所有实验的运行环境为: Windows 10 系统，RTX 3080Ti GPU，AMD Ryzen7 5800X CPU 和 64G 内存。

3.3　基线方法

本文与多种基线方法进行对比，其中包括基于协同过滤的方法、基于图的推荐方法以及跨领域推荐方法。本文设计了两组对比实验，表 3 与表 4 中的单领域表示仅利用目标域的数据训练模型，跨领域表示利用源域和目标域的数据协同训练，其中用户嵌入向量是共享的。所有基线方法的介绍如下：

基于协同过滤的方法　BPRMF[34]是隐因子协同过滤的经典方法，利用用户向量与物品向量的内积作为推荐任务的得分，并优化 BPR 损失函数以最大化正样本得分与负样本得分的差值。

MLP 利用多层感知机(Multilayer Perceptron, MLP)预测用户与物品之间的得分，各个隐层的维度为[40,20,10,1]，激

活函数使用 ReLU。

NeuMF[3]结合矩阵分解与多层感知机，引入用户向量与物品向量的逐元素相乘作为 MLP 的特征，缓解 MLP 难以拟合内积函数的问题。

基于图的推荐方法　NGCF[28]将 GCN 应用于协同过滤模型，在聚合领域信息的过程中，额外考虑了用户与物品之间的相似性。同时，叠加多层图卷积可以提取用户-物品二部图中的高阶领域信息。在对比实验中，本文发现 2 层图卷积的效果最好。

LightGCN[29]是 NGCF 的改进版本，去除每层图卷积的特征变换和非线性层，仅聚合领域特征。与 NGCF 一样，在对比实验中本文堆叠 2 个图卷积层。

跨领域推荐方法　CMF[16] 将两个领域的数据联合起来进行矩阵分解，通过共享的中间变量(用户嵌入矩阵)实现知识迁移的效果。在对比实验中，本文令源领域与目标领域的损失函数权重相等。

DANN[35]是计算机视觉上经典的领域自适应方法，它利用梯度翻转层(Gradient Reverse Layer, GRL)和域分类器学习领域无关的语义特征，本文把它适配于推荐任务，语义特征代表用户的兴趣。采用的骨干网络是 2 层 LightGCN，由于跨领域的是用户，用户表示向量包含领域相关信息，本文在用户表示上添加梯度翻转层与域分类器，同时，再加一个全连接层编码用户的兴趣向量，并利用这个兴趣向量进行推荐。

CD-GNN[33]是近期提出的一种用户跨领域推荐与图神经网络结合的方法，它首先用图神经网络从社交网络上提取用户表示，然后利用一个领域不变层(Domain Invariant Layer)混淆来自不同领域的用户表示，并附加一个域分类器训练域不变层。最后，拼接用户表示与物品表示，用一层全连接网络预测得分。

对于所有基于图的算法，本文构建的输入图包含所有用户与商品节点，但边集仅包含训练集的边。因此，输入图是异质的，它包含用户、源域物品、目标域物品这 3 种节点，以及用户-源域物品、源域物品-用户、用户-目标域物品、目标域物品-用户这 4 种边。

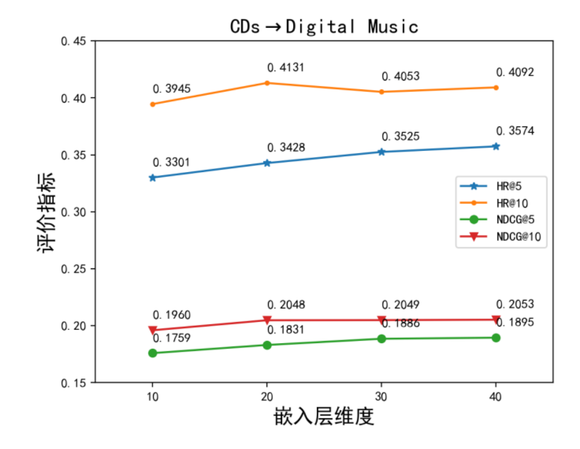
为了进行公平的对比实验，对于所有的基线方法，本文训练最多遍历 1000 次数据集，并采用早停止(early stopping)策略，即在目标领域的验证集误差未下降的 5 次迭代后结束训练，且所有方法均使用 Adam 优化器进行训练，学习率为 1e-3，L2 正则化系数为 1e-4，批次大小为 1024，dropout 概率为 0.2。

3.4　超参数敏感性分析

本节探究本文的模型中几个重要超参数的，包括嵌入层维度大小、BPR 损失函数的负样本比例、以及 L2 正则化系数 λ 。

首先，本文固定了模型的其他超参数，然后采用[10,20,30,40]的嵌入层维度大小进行实验，结果如图 4 所示。在两个任务中，提升嵌入维度至 40 依然可以获得稳定但微弱的提升，与默认维度 20 的差别过小，可以认为模型容量基本上足够了，而提升维度导致计算代价几乎翻倍。因此，本文的模型对嵌入层维度大小并不敏感，在实验中采用的嵌入维度是 20。

其次，本文还尝试不同的 BPR 损失函数的负样本比例，取值范围[1,5,10,20,50,100]，实验结果如图 5 所示。在 CDs→Digital Music 任务上，负样本比例从 1 提升到 50，模型效果有着显著且稳定的提升。比例提升到 100 时，效果略微下降，这可能是大量的负样本中包含了训练集中未观测到的测试样本导致的。在 Movies→Books 任务上，各指标的曲线较为平缓，可能是由目标领域高度稀疏造成的。因此，适当增大负样本比例是有助于模型训练的。在实验中，考虑到性能与计算代价的权衡，本文采用的负样本比例是 10。



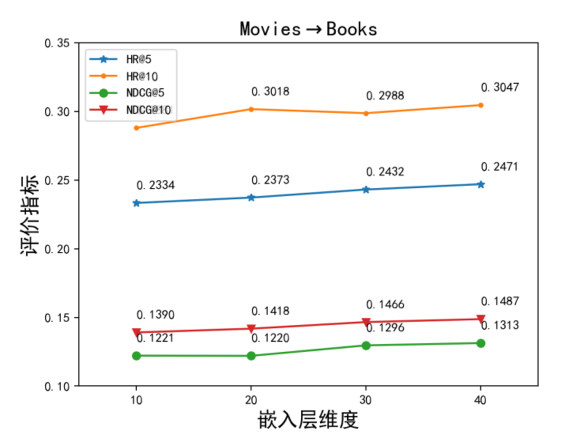
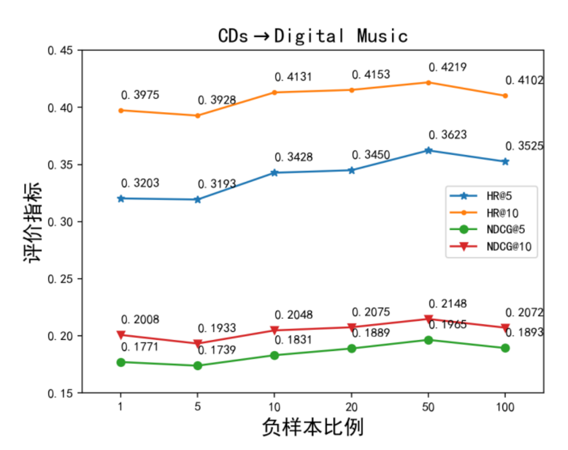


图 4　嵌入层维度的敏感性分析

Fig. 4　Sensitivity analysis of embedding size



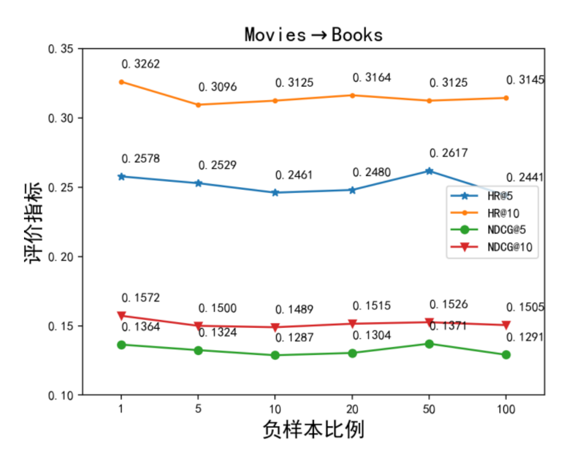
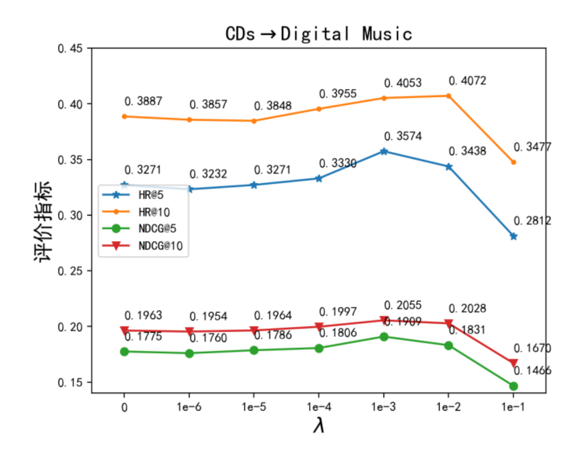


图 5　负样本个数的敏感性分析

Fig. 5　Sensitivity analysis of #negative samples

最后，本文还尝试了不同的 L2 正则化系数，取值范围[0,

1e-6, 1e-5, 1e-4, 1e-3, 1e-2, 1e-1]，实验结果如图 6 所示。在 CDs→Digital Music 任务上，λ =1e-3 取得了最好的效果，而在 Movies→Books 任务上，λ =1e-2 取得了最好的效果。即使在 λ 设置为 0 时，本文的模型对 λ 也相对不敏感。当 λ 大于 1e-2 时，性能下降很快，说明正则化过强会对模型的正常训练产生负面影响。本文采用的 λ 是 1e-3。



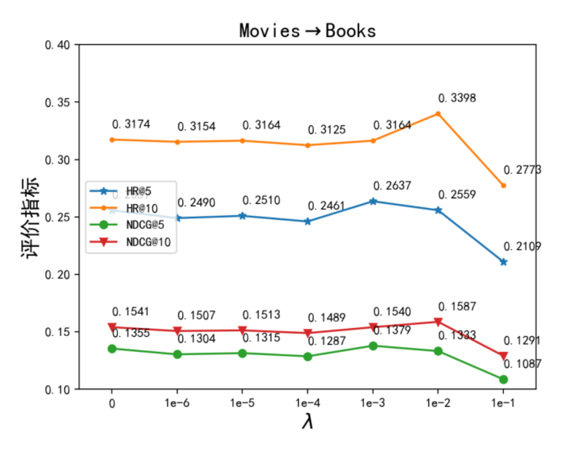


图 6　L2 正则化系数的敏感性分析

Fig. 6　Sensitivity analysis of L2 normalization weight

3.5　实验结果

本文给出实验 CDs→Digital Music 与 Movies→Books 的结果，如表 3 与表 4 所示。对于每组实验，本文分别使用 5 个不同的随机种子并计算平均的实验结果，最好的结果加粗显示。

表 3　实验 CDs→Digital Music 的目标领域测试集结果

Tab. 3　Target test set results on cds→Digital Music

	模型	HR@5	HR@10	NDCG@5	NDCG@10
单领域	MF	0.1668	0.2029	0.0933	0.1044
	MLP	0.1435	0.1994	0.0710	0.0882
	NeuMF	0.1466	0.2118	0.0721	0.0921
	NGCF	0.1802	0.2540	0.0905	0.1131
	LightGCN	0.2746	0.3457	0.1466	0.1685
跨领域	CMF	0.1933	0.2363	0.1059	0.1190
	MLP	0.1506	0.2124	0.0749	0.0938
	NeuMF	0.1506	0.2136	0.0752	0.0945
	NGCF	0.2314	0.3076	0.1192	0.1426
	LightGCN	0.3088	0.3868	0.1631	0.1871
	DANN	0.2116	0.2733	0.1109	0.1298
	CD-GNN	0.1479	0.2144	0.0734	0.0938
	Ours	0.3427	0.4130	0.1830	0.2047
	相对提升	11.0%	6.78%	12.25%	9.44%

与单领域推荐方法相比，本文的方法有着巨大的优势，在所有任务上都获得了最佳的效果。同时，利用了图信息的 NGCF 与 LightGCN 显著优于没有利用图信息的 MF、MLP 和 NeuMF，说明图卷积层提取的用户-物品交互图的结构信息有助于推荐任务。

在跨领域推荐的实验上，本文的方法同样优于所有基线方法。首先，跨领域的方法在大多数结果上优于单领域的方法，与单领域版本相比在 HR@5 上普遍取得 1-2 个百分点的提升，说明了跨领域知识迁移的有效性。其次，在跨领情况下，图卷积方法依旧优于协同过滤方法。此外，DANN 的效果在两个实验中都与 CMF 不相上下，效果远不如其骨干网络 LightGCN，说明了传统的基于计算机视觉的领域自适应技术并不使用于推荐领域。CD-GNN 效果不如 CMF 的原因可能是，原文利用了社交网络的信息，但大多数场景不能获得用户-用户关系，且预测评分的单层全连接并不能建模用户与物品的交互。最后，本文的方法好于其他跨领域推荐方法，与最好的基线方法相比，在 HR@5 上平均提升 14.11%，在 NDCG@5 上平均提升 15.32%，说明了解耦模块能有效地从结构信息中提取用户兴趣。

表 4　实验 Movies→Books 的目标领域测试集结果

Tab. 4　Target test set results on Movies→Books

	模型	HR@5	HR@10	NDCG@5	NDCG@10
单领域	MF	0.1441	0.1911	0.0757	0.0900
	MLP	0.1548	0.2091	0.0798	0.0964
	NeuMF	0.1578	0.2128	0.0812	0.0980
	NGCF	0.1849	0.2503	0.0955	0.1155
	LightGCN	0.2099	0.2799	0.1087	0.1301
跨领域	CMF	0.1636	0.2212	0.0848	0.1025
	MLP	0.1636	0.2200	0.0845	0.1018
	NeuMF	0.1566	0.2149	0.0805	0.0984
	NGCF	0.1978	0.2654	0.1021	0.1229
	LightGCN	0.2057	0.2748	0.1068	0.1280
	DANN	0.1627	0.2213	0.0838	0.1017
	CD-GNN	0.1515	0.2124	0.0770	0.0957
	Ours	0.2460	0.3125	0.1287	0.1489
	相对提升	17.22%	11.64%	18.39%	14.44%

综上，本文的模型显著优于单领域的推荐方法，并相对于现有的跨领域推荐方法有较大提升。

3.6　消融实验

正如前一节所讨论的，本文提出的模型比其他基线方法有了显著的提升。这些改进实际上来自于将两个新组件整合到推荐模型的设计中：解耦模块，它统一了两个领域中用户表示的学习过程，语义编码器提取用户的兴趣，领域编码器提取领域相关的噪声；以及重构模块，结合用户与物品表示重构输入的图，保证了结构信息不会流失。在本节中，本文进行消融实验以证明每个组件的重要性，将提出的模型与以下变体进行比较：

-Gen：在原模型的基础上，删除图重构模块。

-Dom：在原模型的基础上，删除领域编码器与领域分类器。

-Disen：在原模型的基础上，删除整个解耦模块。

如表 5 所示，如果将任何一个模块从推荐模型中移除，都导致 HR 与 NDCG 指标的显著降低。因此，消融实验表明，本文提出的模型的优势实际上来自于这些新模块，这些新模块在双领域协同学习的模型中发挥了重要作用，都有助于提高的跨域推荐性能。

3.7　样例可视化

为了可视化模型的学习结果，本文随机选择 100 个活跃用户(交互行为大于 10)，并查询用户评价过的源领域物品和目标领域物品，再把训练后的用户向量与物品向量拼接，最后用 TSNE 算法投影到 2 维平面。NGCF、LightGCN 与本文的方法在 CDs→Digital Music 数据集上的可视化结果如图 8

7~9 所示。

表 5　消融实验

Tab. 5　Ablation study

数据集	模型	H@5	H@10	N@5	N@10
CDs→Digital Music	Ours	0.3427	0.4130	0.1830	0.2047
	-Gen	0.3369	0.4082	0.1805	0.2025
	-Dom	0.3203	0.3828	0.1760	0.1946
	-Disen	0.3183	0.3789	0.1706	0.1898
Movies→Books	Ours	0.2460	0.3125	0.1287	0.1489
	-Gen	0.2431	0.3115	0.1254	0.1467
	-Dom	0.2343	0.2968	0.1242	0.1430
	-Disen	0.2197	0.2998	0.1160	0.1405

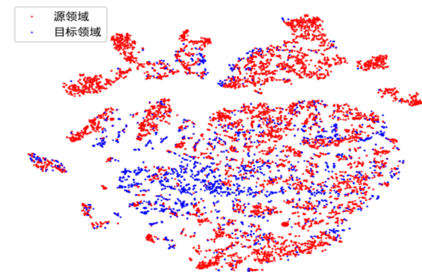


图 7　NGCF 的 TSNE 降维结果

Fig. 7　TSNE projection of NGCF

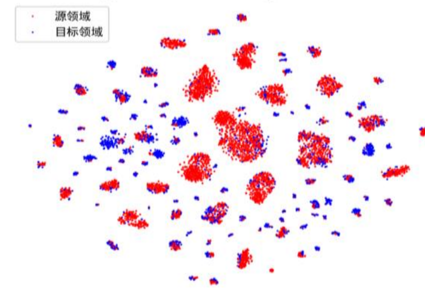


图 8　LightGCN 的 TSNE 降维结果

Fig. 8　TSNE projection of lightgcn

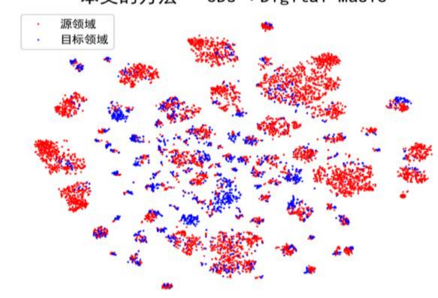


图 9　本文的方法的 TSNE 降维结果

Fig. 9　TSNE projection of our method

在图 7~9 中，红点代表投影后的源领域物品向量，蓝点代表投影后的目标领域物品向量。可以观察到，NGCF 的物品向量投影较为分散，没有形成簇状结构，这表示模型未能捕获物品对应的用户兴趣；LightGCN 的物品向量投影形成了较多的簇，但源领域与目标领域的簇独立存在，即两个域的物品向量未能对齐，有碍于表达跨领域不变的用户兴趣；本文的方法的物品向量投影在呈现簇结构的同时，源领域与目标领域的物品向量充分融合，这表明本文的方法能有效解耦出不随着领域改变的用户兴趣。

4　结束语

本文从图的生成过程角度对跨领域推荐方法进行了研究，本文利用图自编码器把从用户-物品二部图提取的用户表示解耦为多个独立的隐变量，并准确提取了目标用户的跨领域兴趣偏好，提出一种基于图生成过程的跨领域推荐方法。本文在四个真实数据集上进行了大量实验，结果表明本文的方法相对于现有的跨领域推荐方法有较大提升。未来将在图卷积模块尝试更先进的框架，以及考虑多源领域的迁移。

参考文献：

[1] Hu Yifan, Koren Y, Volinsky C. Collaborative filtering for implicit feedback datasets [C]// 2008 Eighth IEEE International Conference on Data Mining. Ieee, 2008: 263-272.

[2] Mnih A, Salakhutdinov R R. Probabilistic matrix factorization [C]// Advances in neural information processing systems. 2008: 1257-1264.

[3] He Xiangnan, Liao Lizi, Zhang Hanwang, *et al.* Neural collaborative filtering [C]// Proceedings of the 26th international conference on world wide web. 2017: 173-182.

[4] Tang Jie, Wu Sen, Ssun Jimeng, *et al.* Cross-domain collaboration recommendation [C]// Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. 2012: 1285-1293.

[5] Singh A P, Gordon G J. Relational learning via collective matrix factorization [C]// Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. 2008: 650-658.

[6] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks [J]. arXiv preprint arXiv: 1609. 02907, 2016.

[7] Hamilton W L, Ying R, Leskovec J. Inductive representation learning on large graphs [C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 1025-1035.

[8] Pazzani M J, Billsus D. Content-based recommendation systems [M]// The adaptive web. Springer, Berlin, Heidelberg, 2007: 325-341.

[9] 张屹晗，王巍，刘华真，等. 基于知识图嵌入的协同过滤推荐算法 [J]. 计算机应用研究, 2021, 38 (12): 3590-3596. (Zhang Yihan, Wang Wei, Li Huazhen, *et al.* Collaborative Filtering Recommendation Algorithm Based on Knowledge Graph Embedding [J]. Application Research of Computers, 2021, 38 (12): 3590-3596.)

[10] 张润莲，张瑞，武小年，等. 基于混合相似度和差分隐私的协同过滤推荐算法 [J]. 计算机应用研究, 2021, 38 (8): 2334-2339. (Zhang Runlian, Zhang Rui, Wu Xiaonian, *et al.* Collaborative Filtering Recommendation Algorithm Based on Mixed Similarity and Differential Privacy [J]. Application Research of Computers, 2021, 38 (8): 2334-2339.)

[11] 王森，陈莉，张洁. 基于项目模糊相似度的协同过滤推荐算法 [J]. 计算机应用研究, 2021 (3): 696-701. (Wang Sen, Chen Li, Zhang Jie. Collaborative Filtering Recommendation Algorithm Based on Item Fuzzy Similarity [J]. Application Research of Computers, 2021 (3): 696-701.)

[12] 苏静，许天琪，张贤坤，等. 基于图卷积与外积的协同过滤推荐模型 [J]. 计算机应用研究, 2021, 38 (10): 5. (Su Jing, Xu Tianqi, Zhang Xianshen, *et al.* Collaborative Filtering Recommendation Model Based on Graph Convolution and Cross Product [J]. Application Research of Computers, 2021, 38 (10): 5.)

[13] Rendle S. Factorization machines [C]// 2010 IEEE International conference on data mining. IEEE, 2010: 995-1000.

[14] 陈燕，马进元，李桃迎. 基于共享评级迁移的跨域推荐算法 [J]. 计算机应用研究, 2021, 38 (9): 6. (Liu Yan, Ma Jinyuan, Li Taoying. Cross Domain Recommendation Algorithm Based on Shared Ratings Transfer [J]. Application Research of Computers, 2021, 38 (9): 6.)

[15] 柴玉梅，员武莲，王黎明，等. 基于双注意力机制和迁移学习的跨领域推荐模型 [J]. 计算机学报, 2020, 43 (10): 19. (Chai Yumei, Yuan

Wulian, Wang Liming, *et al.* A Cross-Domain Recommendation Model Based on Dual Attention Mechanism and Transfer Learning [J]. Chinese Journal of Computers, 2020, 43 (10): 19.)

[16] Singh A P, Gordon G J. Relational learning via collective matrix factorization [C]// Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. 2008: 650-658.

[17] Yuan Feng, Yao Lina, Benatallah B. DARec: Deep domain adaptation for cross-domain recommendation via transferring rating patterns [J]. arXiv preprint arXiv: 1905. 10760, 2019.

[18] Zhu Y, Ge K, Zhuang F, *et al.* Transfer-Meta Framework for Cross-domain Recommendation to Cold-Start Users [C]// Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021: 1813-1817.

[19] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering [J]. Advances in neural information processing systems, 2016, 29: 3844-3852.

[20] Veličković P, Cucurull G, Casanova A, *et al.* Graph attention networks [J]. arXiv preprint arXiv: 1710. 10903, 2017.

[21] Kipf T N, Welling M. Variational graph auto-encoders [J]. arXiv preprint arXiv: 1611. 07308, 2016.

[22] 冯兴杰, 生晓宇. 基于图神经网络与深度学习的商品推荐算法 [J]. 计算机应用研究, 2021, 38 (12): 3617-3622. (Feng Xingjie, Sheng Xiaoyu. Item Recommendation Algorithm Based on GNN and Deep Learning [J]. Application Research of Computers, 2021, 38 (12): 3617-3622.)

[23] 葛尧, 陈松灿. 面向推荐系统的图卷积网络 [J]. 软件学报, 2020, 31 (4): 1101-1112. (Ge Yao, Chen Songcan. Graph Convolutional Network for Recommender Systems [J]. Journal of Software, 2020, 31 (4): 1101-1112.)

[24] 李宇琦, 陈维政, 闫宏飞, 等. 基于网络表示学习的个性化商品推荐 [J]. 计算机学报, 2019, v. 42; No. 440 (08): 1767-1778. (Li Yuqi, Chen Weizheng, Yan Hongfei, *et al.* Learning Graph-Based Embedding for Personalized Product Recommendation [J]. Chinese Journal of Computers, 2019, v. 42; No. 440 (08): 1767-1778.)

[25] 陈劲松, 孟祥武, 纪威宇, 等. 基于多维上下文感知图嵌入模型的兴趣点推荐 [J]. 软件学报, 2020, 31 (12): 3700-3715. (Chen Jingsong, Meng Xiangwu, Ji Weiyu, *et al.* POI Recommendation Based on Multidimensional Context-aware Graph Embedding Model [J]. Journal of Software, 2020, 31 (12): 3700-3715.)

[26] 荣沛, 苏凡军. 基于知识图注意网络的个性化推荐算法 [J]. 计算机应用研究, 2021, 38 (2): 398-402. (Rong Pei, Su Fanjun. Personalized recommendation algorithm based on knowledge graph attention network [J]. Application Research of Computers, 2021, 38 (2): 398-402.)

[27] Berg R, Kipf T N, Welling M. Graph convolutional matrix completion [J]. arXiv preprint arXiv: 1706. 02263, 2017.

[28] Wang Xiang, He Xiangnan, Wang Meng, *et al.* Neural graph collaborative filtering [C]// Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval. 2019: 165-174.

[29] He Xiangnan, Deng Kuan, Wang Xiang, *et al.* Lightgcn: Simplifying and powering graph convolution network for recommendation [C]// Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval. 2020: 639-648.

[30] Wu J, Wang X, Feng F, *et al.* Self-supervised graph learning for recommendation [C]// Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021: 726-735.

[31] Chen H, Wang L, Lin Y, *et al.* Structured graph convolutional networks with stochastic masks for recommender systems [C]// Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021: 614-623.

[32] Zhao Cheng, Li Chenliang, Fu Cong. Cross-domain recommendation via preference propagation graphnet [C]// Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019: 2165-2168.

[33] Liu Ziqi, Shen Yue, Cheng Xiaocheng, *et al.* Learning Representations of Inactive Users: A Cross Domain Approach with Graph Neural Networks [C]// Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 2021: 3278-3282.

[34] Rendle S, Freudenthaler C, Gantner Z, *et al.* BPR: Bayesian personalized ranking from implicit feedback [C]// Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence. 2009: 452-461.

[35] Ganin Y, Ustinova E, Ajakan H, *et al.* Domain-adversarial training of neural networks [J]. The journal of machine learning research, 2016, 17 (1): 2096-2030.